ISSN 1009 - 3095 *Journal of Zhejiang University(SCIENCE)* V.2, No.4, P.406 - 410, Oct. - Dec., 2001
http://www.periodicals.com.cn; http://www.zju.edu.cn;
http://lib.zju.edu.cn/English; http://www.zjupress.com

A BAYESIAN PET RECONSTRUCTION METHOD USING SEGMENTED ANATOMICAL MEMBRANE AS PRIORS*

GONG Tie-zhu(龚铁柱), WANG Yuan-mei(汪元美)

(*The State Key Laboratory of Biomedical Engineering, Zhejiang University, Hangzhou 310027, China*)

Received Oct.16, 2000; revision accepted Mar.21, 2001

Abstract: In this paper a fully Bayesian PET reconstruction method is presented for combining a segmented anatomical membrane a priori. The prior distributions are based on the fact that the radiopharmaceutical activity is similar throughout each region and the anatomical information is obtained from other imaging modalities such as CT or MRI. The prior parameters in prior distribution are considered drawn from hyperpriors for fully Bayesian reconstruction. Dynamic Markov chain Monte Carlo methods are used on the Hoffman brain phantom to gain estimates of the posterior mean. The reconstruction result is compared to those obtained by ML, MAP. Our results showed that the segmented anatomical membrane a priori exhibit improved the noise and resolution properties.

Key words: positron emission tomography, bayesian reconstruction, markov chain monte carlo, segmented anatomical membrane prior

Document code: A **CLC number:** TP391

INTRODUCTION

Positron Emission Tomography (PET) is a powerful biological imaging tool capable of providing high quality functional images. A radioactively labelled substance, designed to show the physiological function of a particular tissue, is injected into a patient. The chosen substance is known to become concentrated in the organ of interest in a manner related to the phenomenon under study. After the radiopharmaceutical is injected, positron emission occurs in the organ at a rate varying spatially according to the concentration. Our purpose is to determine the radiopharmaceutical concentration as a function of position.

The emitted positrons' collision with nearby electrons generate two photons traveling away from each other in opposite directions. The two photons are detected by a pair of opposed detector elements inside an array of discrete detector elements surrounding the patient simultaneous. The PET imaging procedure consists of the collection of a set of number of coincidences by detector pairs. The positron emission is a Poisson process (Vardi et al., 1985) and the photon count is a linear process so that the measured data also have Poisson distribution. Wherein, the variance is equal to the mean, so the noise property of PET is much poorer than MRI or CT.

Statistical methods for image reconstruction can provide improved spatial resolution and noise properties over conventional filtered backprojection (FBP) methods. The maximum likelihood via expectation maximization method (MLEM) is based on the realistic assumption that photon counts follow a Poisson process, and is an appropriate method PET reconstruction (Vardi et al., 1985). However, due to the ill-posed nature of the reconstruction problem, MLEM is known to be unstable and excessively noisy at higher iterations. Some prior constraints can be placed on the estimated image to overcome those questions (Levitan et al., 1987; Mumcuoglu et al., 1994). Most image priors assign low probabilities as a pixel's intensity differs more from its nearest neighbors, those may introduce smoothing across true intensity changes. To alleviate this drawback, in this paper, we propose segmented anatomical membrane prior and assume

* Project supported by NSFC of China(No.69872034) and Zhejiang Provincial Natural Science Foundation.

that the nearby pixels tend to have similar intensities, unless they straddle a border between regions of smoothly varying intensity. The specific locations of these borders can be obtained from high-resolution anatomical images. Dynamic Markov chain Monte Carlo (MCMC) is used to obtain fully Bayesain reconstruction and the parameters in the prior are themselves considered drawn from appropriate hyperpriors. Thus, unlike reconstructions using parameter point estimates such as MAP, the statistic properties, such as variance, of reconstruction results can be easily obtained.

THE SEGMENTED ANATOMICAL MEMBRANE PRIOR

Regardless of the method used for reconstruction, a Bayesian method could be fatally misleading with a bad choice of prior. The fully Bayes formula is as follows:

$$p(\boldsymbol{X}\mid\boldsymbol{Y},\boldsymbol{\theta})\propto p(\boldsymbol{Y}\mid\boldsymbol{X})p(\boldsymbol{X}\mid\boldsymbol{\theta})p(\boldsymbol{\theta}) \quad (1)$$

where $\boldsymbol{X}$, $\boldsymbol{Y}$ are density and measured data, respectively. $p(\boldsymbol{X}\mid\boldsymbol{\theta})$ is the prior density distribution, while $\boldsymbol{\theta}$ is the hyper-parameters vector and $p(\boldsymbol{\theta})$is its prior distribution. $p(\boldsymbol{Y}\mid\boldsymbol{X})$ is the likelihood function.

The prior density of the image is generally specified via Gibbs distributions or Markov Random Fields:

$$p(\boldsymbol{X}\mid\boldsymbol{\theta})=\frac{1}{Z(\boldsymbol{\theta})}\exp(U(\boldsymbol{X},\boldsymbol{\theta})) \quad (2)$$

where $U(\boldsymbol{X},\boldsymbol{\theta})$ is the Gibbs energy function (potential function). $Z(\boldsymbol{\theta})=\sum_{x}U(\boldsymbol{X},\boldsymbol{\theta})$ is the partition function which once assigned, it relates only to the hyper-parameters.

The potential functions of Gibbs distributions are defined on some kinds of neighborhood systems. Most of them assign higher potentials to big intensity differences, which result in global smoothness, such as that of the quadratic potential. These priors can't model larger intensity changes that may occur between different regions. Alternately, some potential functions that increase at a slower rate than quadratic functions, or saturate at a maximum value, have been proposed as means of trading off the desire for local smoothness with the competing requirement that the image not be over smoothed at region boundaries (Mumcuoglu et al., 1994). But none of these methods directly model the presence of image boundaries. The compound Markov random fields (Lee et al., 1995) introduce line process between pixels. The line-site random variables might have two discrete values, 0 and 1, or might be continuously valued from 0 to 1. As choosing the line process type is difficult, the additional computation.

In PET imaging, anatomical tissue type is one of the most influential factors affecting radiopharmaceutical concentration, because different tissue types have different physiological properties. Consequently, a segmentation of the region of interests (ROIs) into different tissue types may strongly resemble its segmentation into different levels of radiopharmaceutical uptake. An anatomical segmentation can be obtained from MRI or CT. However, a significant difference between the PET ROIs and anatomical ones is the partial pixel effect relating to the finite axial and transaxial resolution of PET images. In the reconstruction practice, we also assign some continuity constraints on the density. The prior must ensure the following requirements:

On the interior of each region, the pixels intensities are similar;

On the boundary of different pure regions, the prior must allow interrupted intensity change;

When the boundary is between a pure region and a region of partial pixels, the prior allows a ramp change of image intensity.

Considering the brain imaging, we assume that the tissues can be partitioned into four classes: gray matter (GM), white matter (WM), cerebral spinal fluid (CSF) and partial pixels. The partition information is estimated from some other image segmentation and register process. To modify the partial pixel effect, we first define a selection function for a pixel (i,j) for its nearest neighborhood pixels in potential value calculation:

$$s_{i,j}(x(m,n))=\begin{cases}x(m,n), & l(i,j)=l(m,n)\\ x(i,j), & \text{otherwise}\end{cases} \quad (3)$$

where $x(i,j)$, $l(i,j)$ stand for the density value and the label of pixel (i,j), respectively; the same is the case of $x(m,n)$, $l(m,n)$.

With the selection function, the intensity value at pixel (m, n) is replaced by that at (i, j) when their types are different. But we only use the selection function for the calculation of potential value for those pixels in one region. Otherwise, the intensity value of a pixel keep unchanging. Then we define a modified quadratic prior distribution of pixels in a region,

$$U(\boldsymbol{x}, \beta) = -\beta \sum_{(i,j)\sim(m,n)} [x_{i,j} - s_{i,j}(x(m, n))]^2 \quad (4)$$

where $(i, j) \sim (m, n)$ denotes the 8 neighborhood pixels of (i, j), β, is the hyper-parameter. For the pixels on the boundaries between the pure region or partial region,

$$U(\boldsymbol{x}) = -\sum_{(i,j)\sim(m,n)} \beta (x_{i,j} - x(m, n))^2 \quad (5)$$

Bowsher et al. (1996) introduced a hierarchy prior model using anatomical information. Their method has to segment and reconstruct simultaneously, but does not directly incorporate the anatomical information, and is very complex.

MARKOV CHAIN MONTE CARLO POSTERIOR ESTIMATION

The prevalent method of PET reconstruction is maximum a posteriori estimate (MAP), which is in proper balance between prior expectation and fidelity to the measured data. In the Bayesian paradigm, MAP estimators are the Bayes point estimators for $0-1$ loss function. We used a hierarchical fully Bayesain model as it is difficult for MAP methods to deal with the hyper-parameters problem. At the same time, the point estimator can't be used to obtain the statistical properties of reconstruction results directly.

The Markov chain Monte Carlo simulation method (MCMC) based on the full conditional posterior distribution can resolve these questions (Smith et al., 1993; Wain, 1997). The MCMC method is a fully Bayesian process that we can use to update hyperparameters in each iteration as well as the pixel intensities. As the final estimate is the mean of sampling images, the variance of the reconstruction results and the credit intervals are also obtained.

The basic idea of Markov chain Monte Carlo simulation is to sample from a fully posterior conditional distribution. In our application, the fully posterior conditional distribution is $p(x_{ij} \mid \boldsymbol{X}_{-ij}, \boldsymbol{Y}\beta)$ for pixel intensity, where $\boldsymbol{X}_{-ij} = \{x_{mn}$, except $x_{ij}\}$. Sampling procedures estimate the joint posterior distribution by iterations of sampling for both pixel intensities and hyper-parameters from its conditional distribution given the current values of all other random variables. There are two kinds of sampling algorithms: Gibbs sampler and Metropolis-Hasting sampler. Their updating methods are different. Here we use Metropolis-Hasting sampler.

The Metropolis-Hasting algorithm is as follows: assuming only one random variable, whose posterior distribution is $p(m \mid n)$,, we get the new candidate m' for m from a generation probability distribution, wherein the candidate is accepted with the probability:

$$\min\left\{1, \frac{p(m' \mid n) g(m, m')}{p(m \mid n) g(m', m)}\right\} \quad (6)$$

the second item is called acceptance probability.

In PET images reconstruction, the steps in one iteration can be described below:

1. For each pixel, a candidate intensity value based on the current value is obtained from an appointed generation distribution. With the principle of conjugate prior, we use Gamma distribution $Ga(\hat{x}_{ij}, 1)$ to get the proposed value, where $\hat{x}_{ij}$ denotes current value and Ga for Gamma distribution;

2. Calculate the accepted probability and the update pixel intensity if needed;

3. Update the hyper-parameters after fully updating the image pixels.

We denote the estimator as $\boldsymbol{X} = \{x_{ij}\}$, the measured data $\boldsymbol{Y} = \{y_i\}$, and the system probability array as $\boldsymbol{P} = \{p_{ij}^t\}$. Thqe procedure of a posterior sampling can be performed in two stages. First, we update the pixel values, which we may do according to row or column order. Each time only one pixel is updated while the others are unchanged. Then for the pixel at coordinates (i, j), the acceptance probability is Eq.(7):

$$\frac{p(x'_{ij} \mid \boldsymbol{X}_{-ij}, \boldsymbol{Y}) p(x'_{ij} \mid x_{ij})}{p(x_{ij} \mid \boldsymbol{X}_{-ij}, \boldsymbol{Y}) p(x_{ij} \mid x'_{ij})} = \frac{p(\boldsymbol{Y} \mid x'_{ij}, \boldsymbol{X}_{-ij}) p(x'_{ij}, \boldsymbol{X}_{-ij} \mid \beta) p(\beta)}{p(\boldsymbol{Y} \mid x_{ij}, \boldsymbol{X}_{-ij}) p(x_{ij}, \boldsymbol{X}_{-ij} \mid \beta) p(\beta)} \times \frac{Ga(x_{ij}; x'_{ij}, 1)}{Ga(x'_{ij}; x_{ij}, 1)} =$$

$$\frac{p(\boldsymbol{Y}\mid x'_{ij},\boldsymbol{X}_{-ij})p(x'_{ij},\boldsymbol{X}_{-ij}\mid\beta)}{p(\boldsymbol{Y}\mid x_{ij},\boldsymbol{X}_{-ij})p(x_{ij},\boldsymbol{X}_{-ij}\mid\beta)}\times\frac{Ga(x_{ij};x'_{ij},1)}{Ga(x'_{ij};x_{ij},1)}=\prod_t\exp\{-p_{ij}^t(x'_{ij}-x_{ij})\}\cdot\left(\frac{\sum_{mn}p_{mn}^t x_{mn}+p_{ij}^t(x'_{ij}-x_{ij})}{\sum_{mn}p_{mn}^t x_{mn}}\right)\times\exp\Big(\sum_{ij-mn}U(x_{ij},\boldsymbol{X}_{-ij})-\sum_{ij-mn}U(x'_{ij},\boldsymbol{X}_{-ij})\Big)\times\frac{\Gamma(x_{ij})x_{ij}^{x'_{ij}-1}}{\Gamma(x'_{ij})x'^{x_{ij}-1}_{ij}}\exp(x'_{ij}-x_{ij}) \quad (7)$$

Second, we update the hyper-parameter β. Here we assume that β is drawn from a uniform distribution on an interval centered at the current value and that the lengths of all intervals are equal. The new candidate value is also generated from the same uniform distribution, then the acceptance probability is:

$$\frac{p(\beta'\boldsymbol{X},\boldsymbol{Y})p(\boldsymbol{X}\mid\beta')p(\beta\mid\beta')}{p(\beta'\boldsymbol{X},\boldsymbol{Y})p(\boldsymbol{X}\mid\beta)p(\beta'\mid\beta)}=\frac{p\boldsymbol{Y}\mid\boldsymbol{X}}{p(\boldsymbol{Y}\mid\boldsymbol{X})}\frac{Z(\beta')}{Z(\beta)}\sum_x(\beta-\beta')U(\boldsymbol{X})=\frac{Z(\beta')}{Z(\beta)}\sum_x(\beta-\beta')U(\boldsymbol{X}) \quad (8)$$

To calculate this value, we must know the ratio of partition function at these two points; these ratios can be computed in advance (Weir, 1997).

The last question that must be carefully considered is the detection of convergence. There are many parameters, such as the relative frequency of acceptance of pixels per sweep, several pixel estimates, the hyperparameters, the norm of the increment in image space between successive sweeps, and the total posterior energy that can be monitored. For our practice in the next section, we keep tracing the change of the hyper-parameter.

RESULTS AND DISCUSSION

We performed a reconstruction experiment on one slice of Hoffman head phantom with a size of 128 * 128. The ratio of pharmaceutical concentration in white matter, gray matter and cerebral spinal fluid was 4:1:0. Poisson noise were added to the phantom with a Poisson random variables generation process, c.f. Fig. 1a, b. The segmented prior information was obtained from noise-free data. The project was implemented on 360° with 192 views and 192 projections.

To compare the property of vary methods, we also performed reconstruction with MLEM, MAP

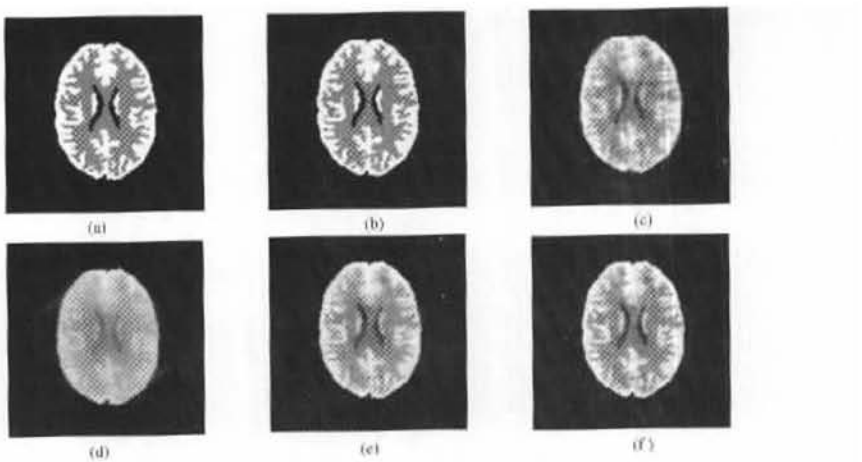

Fig.1 The photon and reconstruction results

(a) hoffman head photon; (b) the Poisson noised image; (c) the result of Ml method; (d) the MAP result with quadratic prior; (e) the MAP result with segmented membrane prior; (f) the end of MCMC with segmented result

with quadratic priori, MAP with the segmented membrane prior distribution (the hyper-parameter was select by the MCMC method) and the MCMC sampler. The four reconstruction results are shown in Fig. 1. The system probability array was precomputed and stored. Both the ML and MAP estimator were calculated by the restarted Fletcher-Reeves conjugate gradient method and the iteration number was limited to 50 times (in general the iteration will converge before 50 times).

For quantitative analysis, we computed the average root mean squared errors (RMSE) of reconstruction images for each method by using the following formula:

$$RMSE = \sqrt{\frac{1}{n_{\text{pixel}}^2} \sum_n (x_n - \lambda_n)^2} \qquad (9)$$

The results are listed below:

Table 1 ***RMSE*** **of each reconstruction methods**

ML	Quadratic prior MAP	MAP of membrane prior	MCMC
13.4	11.28	8.2	6.4

From the computed results, we can see the improvement of our method was phenomenal. The main drawback of MCMC is that it is very time-consuming. In our practice on PC, both the ML and MAP methods converged in several minutes. But the convergence of hyper-parameters in MCMC required several hours; although we expect much faster reconstruction results with the use of more powerful workstations or parallel computations.

References

Bowsher, J. E., Johnson, V. E., Timothy, T. G. et al., 1996. Bayesian reconstruction and use of anatomical a priori information for emission tomography. *IEEE Trans. Med. Imag.*, **15**: 673 – 686.

Lee, S. J., Rangarajan, A., Gindi, G., 1995. Bayesian image reconstruction in SPECT using higher order mechanical models as priors. *IEEE Trans. Med. Imag.*, **14**: 669 – 680.

Levitan, E., Herman, G. T., 1987. A maximum a posteriori probability expectation maximization algorithm for image reconstruction in emission tomography. *IEEE Trans. Med. Imag.*, MI_**6**: 185 – 192.

Mumcuoglu, E., Leachy, R., Cherry, S. R., 1994. Fast gradient-based methods for Bayesian reconstruction of transmission and emission PET images. *IEEE Trans. Med. Imag.*, **13**: 687 – 701.

Smith, A. F. M., Roberts, G. O., 1993. Bayesian computation via Gibbs sampler and related Markov chain Monte Carlo methods. *J. R. Statist. Soc. B*, **55**: 3 – 23.

Vardi, Y., Shepp, L. A., Kaufman, L., 1985. A statistical model for positron emission tomography. *J. Amer. Statist. Ass.*, **80**: 8 – 20.

Weir, I. S., 1997. Fully Bayesian reconstructions from single-photon emission computed tomography data. *J. Amer. Statistic. Ass.*, **92**: 49 – 60.